Groupe de travail de théorie ergodique
Invariance de l'entropie par équivalence orbitale de Shannon pour les actions p.m.p. de $\mathbb{Z}$

Corentin Correia

13 juin 2023

Table des matières

1 Introduction

Nous allons présenter la preuve du théorème suivant, dû à David Kerr et Hanfeng Li dans [6].

Théorème 1. *Soit $T \in \mathrm{Aut}(X,\mu)$ et $U \in \mathrm{Aut}(Y,\nu)$ deux transformations apériodiques et ergodiques. Si elles sont orbitalement équivalentes et que parmi les deux cocycles associés à l'équivalence orbitale, $c_T\colon X \mapsto \mathbb{Z}$ est Shannon, alors $h(T) \leqslant h(U)$.*

Les espaces de probabilité (X,μ) et (Y,ν) sont toujours supposés standards. Les deux auteurs donnent en fait un énoncé plus général qui est le suivant.

Théorème 2. *Soit G et H deux groupes infinis, finiment engendrés. On suppose que G est virtuellement cyclique, H est virtuellement abélien et que $G \overset{\alpha}{\curvearrowright} (X,\mu)$ et $H \overset{\beta}{\curvearrowright} (Y,\nu)$ sont deux actions p.m.p. libres et orbitalement équivalentes et que parmi les deux cocycles associés à l'équivalence orbitale, $\kappa_G\colon G \times X \to H$ est Shannon. Alors $h(\alpha) \leqslant h(\beta)$.*

On dit qu'un groupe vérifie virtuellement une propriété si cette propriété est vérifiée par un sous-groupe d'indice fini (par exemple : virtuellement cyclique, virtuellement abélien).

De cette énoncé asymétrique par rapport aux cocycles, on en déduit une version symétrique : si on suppose de plus que H est aussi virtuellement cyclique et que l'autre cocycle $\kappa_H\colon H \times Y \to G$ est Shannon (dans ce cas on parle d'équivalence orbitale de Shannon), alors $h(\alpha) = h(\beta)$. Kerr et Li en déduisent un résultat plus général, s'appuyant sur leurs travaux dans [5] :

Théorème 3. *Soit G et H deux groupes infinis, finiment engendrés et virtuellement abéliens. Si $G \overset{\alpha}{\curvearrowright} (X, \mu)$ et $H \overset{\beta}{\curvearrowright} (Y, \nu)$ sont deux actions p.m.p. et libres, Shannon orbitalement équivalentes, alors elles ont même entropie : $h(\alpha) = h(\beta)$.*

Nous nous restreignons au cas plus simple d'actions du groupe $\mathbb{Z}$ (c'est-à-dire $G = H = \mathbb{Z}$). Par conséquent, nous parlerons de transformations p.m.p. $T\colon X \to X$ et $U\colon Y \to Y$ au lieu d'actions de groupes p.m.p., de transformations apériodiques au lieu d'actions libres. Avec cette simplification de l'énoncé, les preuves dans [6] sont raccourcies, nous expliquerons en note de bas de page les étapes supplémentaires que nécessite le cas général dans les preuves. Nous ne pourrons cependant pas nous passer de certains résultats généraux sur les groupes moyennables, notamment pour les calculs d'entropie, ce que nous expliquerons dans la partie 2.4.

Les preuves proposées dans [6] nécessitent l'ergodicité des deux actions mais un argument de Kerr et Li (utilisant le théorème de décomposition ergodique et la formule intégrale de l'entropie par rapport à cette décomposition, voir [1], [2]) permet de généraliser le résultat obtenu aux actions non nécessairement ergodiques. Ici, nous supposerons de toute façon l'ergodicité dès qu'on en a besoin et nous garderons en tête que le résultat reste vrai sans ergodicité.

2 Prérequis : définitions, notations, résultats importants

2.1 Equivalence orbitale

Définition 1. *Deux transformations apériodiques $T \in \mathrm{Aut}(X, \mu)$ et $U \in \mathrm{Aut}(Y, \nu)$ sont dites orbitalement équivalentes s'il existe des parties X_0 de X et Y_0 de Y de mesure pleine, respectivement T- et U-invariante, et φ un isomorphisme mesuré de X_0 vers Y_0 tels que T et $\varphi^{-1}U\varphi$ ont les mêmes orbites sur X_0. φ est appelée une équivalence orbitale entre T et U.*

Dans ce cas, on peut définir presque partout et de manière unique (par apériodicité) des fonctions $c_T\colon X \mapsto \mathbb{Z}$ et $c_U\colon Y \mapsto \mathbb{Z}$ par :

$$Tx = \varphi^{-1}U^{c_T(x)}\varphi(x) \text{ et } Uy = \varphi T^{c_U(y)}\varphi^{-1}(y),$$

c_T et c_U sont appelés les cocycles associés à cette équivalence orbitale. Plus généralement, on définira presque partout des fonctions $\kappa_T\colon \mathbb{Z} \times X \mapsto \mathbb{Z}$ et $\kappa_U\colon \mathbb{Z} \times Y \mapsto \mathbb{Z}$ par :

$$T^n x = \varphi^{-1}U^{\kappa_T(n,x)}\varphi(x) \text{ et } U^n y = \varphi T^{\kappa_U(n,x)}\varphi^{-1}(y),$$

qu'on appellera également des cocycles. A noter que $\kappa_T(1, x) = c_T(x)$ et $\kappa_U(1, x) = c_U(x)$. [1]

Dans la suite, nous supposerons sans perte de généralité que $(X, \mu) = (Y, \nu)$ et que l'identité sur X est une équivalence orbitale.

Propriété 1. *Par exemple pour κ_T, on a l'identité de cocycle :*

$$\forall m, n \in \mathbb{Z},\ \kappa_T(m + n, x) = \kappa_T(m, x) + \kappa_T(n, T^m x),$$ [2]

de laquelle on en déduit :

$$\forall j, n \in \mathbb{Z},\ \kappa_T(jn, x) = \kappa_T(j, x) + \kappa_T(j, T^j x) + \ldots + \kappa_T(j, T^{j(n-2)}x) + \kappa_T(j, T^{j(n-1)}x). \tag{1}$$

Par exemple, avec $j = 1$, on obtient :

$$\kappa_T(n, x) = c_T(x) + c_T(Tx) + \ldots + c_T(T^{n-2}x) + c_T(T^{n-1}x).$$

1. Ici on profite du fait que $\mathbb{Z}$ est engendré par un élement. Dans le cas général de groupes agissant sur X, non nécessairement cycliques, on ne peut pas définir des fonctions analogues c_G et c_H, on ne dispose que des cocycles $\kappa_G\colon G \times X \to H$ et $\kappa_H\colon H \times Y \to G$ définis par $gx = \varphi^{-1}(\ \kappa_G(g, x)\varphi(x)\ )$ et $hy = \varphi(\ \kappa_H(h, y)\varphi^{-1}(y)\ )$.
2. Dans le cas général, $\kappa_G(g_1g_2, x) = \kappa_G(g_1, g_2x)\kappa_G(g_2, x)$.

2.2 Quelques notations sur les partitions

Etant donnée une partition $\mathcal{P}$ de X, une transformation $T \in \mathrm{Aut}(X, \mu)$ et F une partie finie de $\mathbb{Z}$, $\mathcal{P}^F$ désigne la partition jointe $\bigvee_{m\in F} T^{-m}(\mathcal{P})$. Dans le cas particulier $F = [\![0, n-1]\!]$, on notera $\mathcal{P}^n$. Mêmes notations pour un recouvrement ouvert fini $\mathcal{U}$ d'un espace topologique.

On définit souvent une partition $\mathcal{P}$ à partir d'une observable, c'est-à-dire une certaine application mesurable $f\colon X \mapsto I$ surjective avec I dénombrable discret (très souvent $I \subset \mathbb{Z}$), en posant $\mathcal{P} := (f^{-1}(n))_{n\in I}$. Chaque pièce est entièrement caractérisée par les valeurs prises par l'observable dessus. Les partitions qu'on définit de cette manière codent une certaine information qui peut devenir de plus en plus compliquée à visualiser au fur et à mesure des différentes opérations sur celles-ci (par exemple lorsqu'on joint de telles partitions). Il nous faut une notation donnant un moyen plus visuel de comprendre l'information codée.

Une partition $\mathcal{P}$ définit elle-même une observable $\mathcal{P}\colon x \in X \mapsto \mathcal{P}(x) \in \mathcal{P}$ où $\mathcal{P}(x)$ est la pièce qui contient x. On utilisera donc le point de vue plus pratique consistant à confondre observable f et partition $\mathcal{P}$ associée. On écrira par abus $\mathcal{P} = f$, on se permettra de noter $H(f)$ au lieu de $H(\mathcal{P})$, qui vaut donc $\sum_{n\in I} -\mu(f^{-1}(n)) \log \mu(f^{-1}(n))$, "$P \in f$" signifiera "$P \in \mathcal{P}$", etc.

On vérifie les propriétés suivantes vis-à-vis des opérations sur les partitions :

- Si f et g sont des observables de $\mathcal{P}$ et $\mathcal{Q}$ respectivement, alors (f, g) est une observable de $\mathcal{P} \vee \mathcal{Q}$, qu'on pourra également noter $f \vee g$ par abus ;
- $f \circ T^m$ est une observable de $T^{-m}(\mathcal{P})$. Ainsi $\mathcal{P}^F$ est d'observable $(f \circ T^m)_{m\in F}$;
- Si C est une partie mesurable de X, alors $f_{|C}\colon C \to f(C)$ est une observable de la partition $\mathcal{P}_C := \{P \cap C \mid P \in \mathcal{P},\ P \cap C \neq \varnothing\}$ de C.

Avec ces exemples, on constate l'importance de la notation pour ne pas perdre de vue l'information qui est codée par une partition, par exemple pour la partition jointe $\mathcal{P}^F = (f \circ T^m)_{m\in F}$: on identifie chaque pièce de ce qu'on observe de chaque $f \circ T^m$, $m \in F$. Un autre intérêt est de lever l'ambiguité de la notation $\mathcal{P}^F$ lorsqu'on dispose de deux transformations bimesurables T et U : est-ce $\bigvee_{m\in F} T^{-m}(\mathcal{P})$ ou $\bigvee_{m\in F} U^{-m}(\mathcal{P})$? On pourra noter à la place $(\mathcal{P} \circ T^m)_{m\in F}$ ou $(\mathcal{P} \circ U^m)_{m\in F}$, ce sera utile dans la preuve du lemme 1.

Le cardinal $|f|$ d'une observable sera définie comme le cardinal de son image, ce qui correspond au cardinal de la partition associée par surjectivité.

Pour finir, on suppose connues les propriétés de l'entropie d'une partition (monotonie par rapport à la relation de finesse, sous-additivité pour des partitions jointes, majoration par le logarithme du cardinal de la partition, etc). Etant donné que μ sera très souvent la mesure de probabilité considérée, on notera H au lieu de H_μ, h au lieu de h_μ.

2.3 Partitions associées aux cocycles, équivalence orbitale de Shannon

Définition 2. *Les cocycles κ_T et κ_U fournissent des partitions $\mathcal{Q}_n^T = \{X_{n,m}^T \mid m \in \mathbb{Z}\}$ et $\mathcal{Q}_n^U = \{X_{n,m}^U \mid m \in \mathbb{Z}\}$ de X avec $X_{n,m}^T = \{x \in X \mid \kappa_T(n, x) = m\}$ et $X_{n,m}^U = \{x \in X \mid \kappa_U(n, x) = m\}$. Par définition, on a $\mathcal{Q}_n^T = \kappa_T(n, \cdot)$ et $\mathcal{Q}_n^U = \kappa_U(n, \cdot)$.*

Propriété 2. *En utilisant (1) dans le cas $j = 1$, on constate que la partition $(\mathcal{Q}_1^T)^n = \left(c_T \circ T^i\right)_{0\leqslant i\leqslant n-1}$ raffine $\mathcal{Q}_{n-1}^T$. Ainsi les $\mathcal{Q}_n^T$ sont toutes d'entropie finie si et seulement si $\mathcal{Q}_1^T$ est d'entropie finie.* [3]

Définition 3. *On dit que c_T est Shannon si $H(c_T) = H(\mathcal{Q}_1^T)$ est finie. D'après la propriété précédente, c'est équivalent à demander que pour tout n, l'entropie $H(\kappa_T(n, \cdot)) = H(\mathcal{Q}_n^T)$ soit finie.* [4]

On dit de deux transformations orbitalement équivalentes qu'elles sont Shannon orbitalement équivalentes si les deux cocycles associés sont Shannon. [5]

3. Ceci est dû au fait que $\mathbb{Z}$ est engendré par 1. Dans le cas général de groupes agissant sur X, pour que toutes les partitions soient d'entropie finie, il suffit que ce soit le cas pour les partitions associées à des éléments d'une partie génératrice du groupe.

4. Dans le cas général de groupes agissant sur X, on dira que le cocycle $\kappa_G\colon G \times X \to H$ est Shannon si toutes les partitions sont d'entropie finie, ou au moins celles associées à des éléments d'un sous-ensemble générateur.

5. Comme annoncé dans l'introduction, on ne s'intéressera qu'au cas où l'un des deux cocycles est Shannon.

2.4 Suites de Følner, entropie mesurée et entropie topologique de transformations, théorème ergodique

On suppose connues :

1. l'entropie (mesurée) d'une transformation T d'un espace mesuré (X, μ) par rapport à une partition $\mathcal{P}$: $h(T, \mathcal{P}) := \lim_{n\to+\infty} \frac{1}{n} H(\mathcal{P}^n)$;
2. l'entropie topologique de T d'un espace topologique $(X, \mathcal{O})$ par rapport à un recouvrement ouvert finie $\mathcal{U}$: $h_{top}(T, \mathcal{U}) := \lim_{n\to+\infty} \frac{1}{n} \log N(\mathcal{U}^n)$ où N compte le cardinal minimal d'un sous-recouvrement ;
3. l'entropie de T : $h(T) := \sup_{\mathcal{P}} h(T, \mathcal{P})$;
4. l'entropie topologique de T : $h_{top}(T) := \sup_{\mathcal{U}} h_{top}(T, \mathcal{U})$;
5. les théorèmes suivants : les deux dernières bornes supérieures sont atteintes respectivement pour une partition $\mathcal{P}$ génératrice (théorème de Kolmogorov-Sinaï) et pour un recouvrement ouvert fini $\mathcal{U}$ générateur ;
6. le principe variationnel : $h_{top}(T) = \sup_{\mu} h(T)$ où la borne supérieure porte sur toutes les mesures de probabilité sur X préservées par T.

On a cependant besoin de quelques outils supplémentaires pour manipuler ces quantités, qui proviennent d'un cadre plus général : l'entropie (mesuré ou topologique) d'une action d'un groupe moyennable. On va se servir du fait que $\mathbb{Z}$ est un groupe infini dénombrable qui est de plus moyennable (i.e. admettant une suite de Følner, cf définition 4) pour obtenir d'autres moyens de calculer les entropies. Les suites de Følner permettent de généraliser le rôle de $[\![0, n]\!]$. Ces résultats sont vrais en remplaçant $\mathbb{Z}$ par un groupe moyennable quelconque, on peut les retrouver dans [7].

Définition 4. *On dit qu'une suite $(F_n)_{n\in\mathbb{N}}$ de parties finies de $\mathbb{Z}$ est de Følner si pour tout $k \in \mathbb{Z}$,*

$$\frac{1}{|F_n|}|(k + F_n)\Delta F_n| \underset{n\to+\infty}{\longrightarrow} 0.$$

Etant donnés F et K des parties finies de $\mathbb{Z}$ et $\delta > 0$, on dit que F est (K, δ)-invariante si

$$|\{m \in F \mid K + m \subset F\}| \geqslant (1 - \delta)|F|.$$

Exemple 1. *$([\![0, n]\!])_{n\in\mathbb{N}}$ et $([\![-n, n]\!])_{n\in\mathbb{N}}$ sont de Følner dans $\mathbb{Z}$.*

Proposition 1.

- *Soit $(F_n)_{n\in\mathbb{N}}$ une suite de parties finies de $\mathbb{Z}$. Il y a équivalence entre :*
  - *i) $(F_n)_{n\in\mathbb{N}}$ est une suite de Følner ;*
  - *ii) pour toute partie finie K de $\mathbb{Z}$ et pour tout $\delta > 0$, il existe $N \in \mathbb{N}$ tel que F_n est (K, δ)-invariante pour tout $n \geqslant N$.*

  ii) est équivalent au même énoncé mais en se restreignant aux parties finies K incluses dans une partie génératrice du groupe. Pour $\mathbb{Z}$, le problème se réduit à $K = \{1\}$.

- *Soit $f \colon \mathcal{P}_f(\mathbb{Z}) \to \mathbb{R}$ et $L \in \mathbb{R}$. Il y a équivalence entre :*
  - *i) Pour toute suite de Følner $(F_n)_{n\in\mathbb{N}}$, $f(F_n) \underset{n\to+\infty}{\longrightarrow} L$;*
  - *ii) pour tout $\varepsilon > 0$, il existe une partie finie K de $\mathbb{Z}$ et $\delta > 0$ tels que $|f(F) - L| \leqslant \varepsilon$ pour toute partie finie F qui est (K, δ)-invariante.*

  Dans ce cas, on dit que $f(F)$ converge vers L lorsque F devient de plus en plus invariante, noté $f(F) \xrightarrow[F\to+\infty]{inv} L$ ou $L = \overset{inv}{\lim}_{F\to+\infty} f(F)$.

  De même, on peut se restreindre aux K inclus dans une partie finie génératrice du groupe, ce qui rend les choses beaucoup plus simples pour $\mathbb{Z}$ car K est fixé, égal à $\{1\}$, et il n'y a plus qu'à prouver l'existence d'un δ.

La proposition précédente met en évidence un nouveau mode de convergence dont les suites de Følner incarnent une caractérisation séquentielle. D'après le théorème qui suit, c'est en fait ce mode de convergence qui a lieu pour définir l'entropie d'une transformation par rapport à une partition ou un recouvrement, on généralise donc les énoncés prérequis au debut de cette sous-partie.

Théorème 4. • $\frac{1}{|F|}H(\mathcal{P}^F) \xrightarrow[F\to+\infty]{inv} h(T,\mathcal{P})$.

De plus, $h(T,\mathcal{P}) = \inf_{F\in\mathcal{P}_f(\mathbb{Z})} \frac{1}{|F|}H(\mathcal{P}^F) = \inf_{n\in\mathbb{N}} \frac{1}{|F_n|}H(\mathcal{P}^{F_n})$ *pour toute suite* (F_n) *de Følner.*

- $\frac{1}{|F|}\log N(\mathcal{U}^F) \xrightarrow[F\to+\infty]{inv} h_{top}(T,\mathcal{U})$.

Les suites de Følner jouent ici le même rôle que $(\llbracket 0,n\rrbracket)_{n\in\mathbb{N}}$, ce phénomène ne s'arrête pas aux calculs d'entropie car il a également lieu dans le théorème ergodique de von Neumann.

Théorème 5 (Théorème ergodique de von Neumann, cas ergodique)**.** *Soit* $T \in \mathrm{Aut}(X,\mu)$ *ergodique. Pour tout* $f \in \mathrm{L}^2(\mu)$*, pour toute suite de Følner* (F_n)*,* $\left(\frac{1}{|F_n|}\sum_{i\in F_n} f\circ T^i\right)_{n\geqslant 1}$ *tend vers* $\int_X f\mathrm{d}\mu$ *dans* $\mathrm{L}^2(\mu)$*.*

On peut également se demander si le théorème ergodique ponctuel de Birkhoff s'énonce avec des suites de Følner. Ceci est vrai avec une hypothèse supplémentaire sur la suite (voir [7]).

2.5 Partitions U-uniformes

On a besoin en quelque sorte de faire un lien entre l'entropie mesurée d'une transformation et l'entropie topologique d'un certain facteur. En effet, dans la preuve du lemme 1 énoncé dans la suite, on devra faire apparaître l'entropie de U dans un majorant de l'entropie de T. Or une entropie d'une partition peut se majorer par le logarithme de son cardinal, une quantité qui ressemble un peu à la définition de l'entropie topologique. La notion de partition U-uniforme va beaucoup nous aider (voir [2], [3], [4], [8], [9]).

Etant donnée $U \in \mathrm{Aut}(X,\mu)$ et une partition $\mathcal{P}$, on définit l'application de codage $q_{\mathcal{P}} \colon x \in X \mapsto (\mathcal{P}(U^i x))_{i\in\mathbb{Z}} \in \mathcal{P}^{\mathbb{Z}}$, ainsi que $E_{\mathcal{P}}$ le support de $\nu_{\mathcal{P}} := (q_{\mathcal{P}})_*\mu$ (⚠ $\mathcal{P}^{\mathbb{Z}}$ désigne le produit cartésien, aucun rapport avec les partitions jointes). $E_{\mathcal{P}}$ est fermé et invariant par le décalage $D_{\mathcal{P}} \colon \mathcal{P}^{\mathbb{Z}} \to \mathcal{P}^{\mathbb{Z}}$. $(D_{\mathcal{P}}, E_{\mathcal{P}}, \nu_{\mathcal{P}})$ est un facteur de (U,X,μ) via l'application équivariante $q_{\mathcal{P}}$, et donc $h_\mu(U) \geqslant h_{\nu_{\mathcal{P}}}(D_{\mathcal{P}})$.

Si U est ergodique, on dit qu'une partition finie $\mathcal{P}$ est U-uniforme si la convergence dans le théorème ergodique ponctuel appliqué à chaque indicatrice des pièces de $\mathcal{P}$ est uniforme sur une partie de X de mesure pleine. C'est équivalent au fait que le système $(D_{\mathcal{P}}, E_{\mathcal{P}})$ est uniquement ergodique. Ainsi dans le cas d'une partition finie $\mathcal{P}$ qui est U-uniforme, on a $h_{\nu_{\mathcal{P}}}(D_{\mathcal{P}}) = h_{top}(D_{\mathcal{P}})$ d'après le principe variationnel. Or dans $\mathcal{P}^{\mathbb{Z}}$, la partition donnée par les cylindres est un recouvrement ouvert fini qui est topologiquement générateur donc $\frac{1}{|F|}\log\left|\left(\mathcal{P}\circ U^i\right)_{i\in F}\right| \xrightarrow[F\to+\infty]{\mathrm{inv}} h_{top}(D_{\mathcal{P}})$ (⚠ on n'écrira pas $\mathcal{P}^F$, on n'utilisera cette notation seulement pour l'action de T).

Au total, on obtient que pour une partition finie $\mathcal{P}$ qui est U-uniforme, $\overset{\mathrm{inv}}{\lim}_{F\to+\infty} \frac{1}{|F|}\log\left|\left(\mathcal{P}\circ U^i\right)_{i\in F}\right|$ existe et est plus petite que $h_\mu(U)$.

Autre propriétés importantes qui nous seront utiles :

Les partitions finies U-uniformes sont denses dans l'ensemble des partitions finies muni de la distance de Rokhlin $(\mathcal{P},\mathcal{Q}) \mapsto H_\mu(\mathcal{P}\mid\mathcal{Q}) + H_\mu(\mathcal{Q}\mid\mathcal{P})$. L'application $\mathcal{P} \mapsto H_\mu(T,\mathcal{P})$ étant continue sur l'ensemble des partitions finies muni de cette distance, on peut donc se restreindre aux partitions finies U-uniformes.

Enfin, si $\mathcal{P}$ est une partition U-uniforme et R une partie finie de $\mathbb{Z}$, alors $\mathcal{P}^R$ (partition jointe pour l'action de T) est U-uniforme.

3 Démonstration du théorème 1

On définit
$$\tilde{h}(\kappa_T) := \inf_{k\in\mathbb{N}^*} \frac{1}{k} H(\mathcal{Q}_k^T)$$ [6]

Le but est de démontrer les lemmes suivants. On en déduira le théorème 1.

Lemme 1. *Soient $T, U \in \mathrm{Aut}(X,\mu)$ deux transformations apériodiques et ergodiques. On suppose qu'elles ont les mêmes orbites et que c_T est Shannon. Alors $h(T) \leqslant \tilde{h}(\kappa_T) + h(U)$.*

Lemme 2. *Soient $T, U \in \mathrm{Aut}(X,\mu)$ deux transformations apériodiques et ergodiques. On suppose qu'elles ont les mêmes orbites et que c_T est Shannon. Alors $\tilde{h}(\kappa_T) = 0$.*

3.1 Un lemme intermédiaire

Le lemme suivant permet de démontrer le lemme 1.

Lemme 3. *Soient $T, U \in \mathrm{Aut}(X,\mu)$ deux transformations apériodiques et ergodiques. On suppose qu'elles ont les mêmes orbites. Soit $(F_n)_{n\in\mathbb{N}}$ une suite de Følner de $\mathbb{Z}$ et $\delta > 0$. Alors pour n assez grand, on dispose d'un ensemble mesurable X_n de X tel que :*

1. *$\mu(X_n) \geqslant 1 - \delta$;*
2. *$\forall x \in X_n,\ \kappa_T(F_n, x)$ est $(\{1\}, \delta)$-invariante.*

On se servira de ce lemme pour approcher $\overset{\text{inv}}{\lim}_{F\to\infty} \frac{1}{|F|} \log |\mathcal{P}^F|$. Nous donnons deux preuves de ce lemme, la première est une adaptation de la preuve de [6] dans le cas du groupe $\mathbb{Z}$, la seconde est une autre approche qui utilise l'ergodicité de T pour appliquer le théorème ergodique de von Neumann.

Preuve du lemme 3, adaptée de celle dans [6].

- Par apériodicité, on obtient l'injectivité de $i \mapsto \kappa_T(i,x)$ et des propriétés comme "$T^A x \subset T^B x \iff A \subset B$" par exemple (idem pour U). On obtient alors :

$$\begin{aligned}
|\{j \in \kappa_T(F_n,x) \mid 1 + j \in \kappa_T(F_n,x)\}| &= |\{i \in F_n \mid 1 + \kappa_T(i,x) \in \kappa_T(F_n,x)\}| \\
&= |\{i \in F_n \mid U^{1+\kappa_T(i,x)} \in U^{\kappa_T(F_n,x)}\}| \\
&= |\{i \in F_n \mid UT^i x \in T^{F_n} x\}| \\
&= |\{i \in F_n \mid T^{c_U(T^i x)+i} x \in T^{F_n} x\}| \\
&= |\{i \in F_n \mid c_U(T^i x) + i \in F_n\}|
\end{aligned}$$

  On veut que $|\{i \in F_n \mid c_U(T^i x) + i \in F_n\}|$ soit plus grand que $(1-\delta)|F_n|$. Après avoir remarqué que $|F_n| = |\kappa(F_n,x)|$, on obtiendra ainsi que $\kappa_T(F_n,x)$ est $(\{1\},\delta)$-invariante et il faudra aussi montrer que c'est vrai pour x appartenant à une partie de mesure au moins $1-\delta$.

- Soit V un sous-ensemble de X, union fini d'éléments de la partition de c_U. Si $T^i x \in V$ pour tout $i \in F_n$, alors on a le résultat voulu. En effet, $c_U(V)$ est fini et (F_n) est de Følner donc (F_n) est $(c_U(V),\delta)$-invariante pour n assez grand. Cependant, il n'y a aucune raison pour que tous les $T^i x$ soit dans V, notons $F'_n(x) = \{i \in F_n \mid T^i x \in V\}$ (en prenant V de plus en plus grand, $F_n \backslash F'_n(x)$ mesurera l'erreur qui est faite en approximant X par V). Alors

$$\begin{aligned}
|\{i \in F_n \mid c_U(T^i x) + i \subset F_n\}| &\geqslant |\{i \in F'_n(x) \mid c_U(V) + i \subset F_n\}| \\
&\geqslant |\{i \in F_n \mid c_U(V) + i \subset F_n\}| - |F_n \backslash F'_n(x)|.
\end{aligned}$$

  $c_U(V)$ est fini donc F_n est $(c_U(V),\delta/2)$-invariante pour n assez grand, ce qui donne au total :

$$|\{j \in \kappa_T(F_n,x) \mid 1 + j \in \kappa_T(F_n,x)\}| \geqslant (1 - \frac{\delta}{2})|F_n| - |F_n \backslash F'_n(x)|.$$

6. Dans [6], on définit cette quantité dans le cas général d'un groupe G virtuellement cyclique. Il s'agit de $\inf_g \frac{1}{|G:\langle g\rangle|} H(\mathcal{Q}_g^G)$ où la borne inférieure porte sur l'ensemble des $g \in G$ qui engendrent un sous-groupe distingué et d'indice fini de G. Pour de tels éléments g, on montre que cela vaut aussi $\inf_{k\in\mathbb{N}^*} \frac{1}{|G:\langle g^k\rangle|} H(\mathcal{Q}_{g^k}^G)$.

- Il reste à choisir V de sorte que $X_n := \{x \in X \mid |F_n \backslash F'_n(x)| \leqslant (\delta/2)|F_n|\}$ soit de mesure plus grande que $1-\delta$. Notons $f_n(x) := |F_n \backslash F'_n(x)|/|F_n(x)| = (1/|F_n|)\sum_{i\in F_n} \mathbb{1}_{T^{-i}(X\backslash V)}(x)$. Alors
$$\mu(X_n) = \mu(f_n \leqslant \delta/2) \geqslant 1 - \frac{\int_X f_n}{\delta/2} = 1 - \frac{\mu(X\backslash V)}{\delta/2},$$
la dernière égalité venant du théorème de Fubini-Tonelli et du fait que T est préserve μ. Prenons donc V assez grand de sorte que $\mu(V) \geqslant 1 - \delta^2/2$, ce qui est possible car la partition de c_U est dénombrable. Avec un tel V, on obtient $\mu(X_n) \geqslant 1-\delta$, ce qui conclut la preuve.

□

Preuve revisitée du lemme 3. [7]

La preuve commence de la même manière que la précédente, on constate que :
$$|\{j \in \kappa_T(F_n,x) \mid 1+j \notin \kappa_T(F_n,x)\}| = |\{i \in F_n \mid c_U(T^i x) + i \notin F_n\}|.$$

Si K est une partie finie de $\mathbb{Z}$, alors pour n assez grand, F_n est $(K,\delta/2)$ invariante, ce qui signifie que $G_n := \{i \in F_n \mid K + i \subset F_n\}$ est de cardinal $\geqslant (1-\delta/2)|F_n|$ et vérifie $K + G_n \subset F_n$. Ainsi si $c_U(T^i x)+i$ n'est pas dans F_n alors que i est dans G_n, c'est que $c_U(T^i x)$ n'est pas dans K. En comptant également les autres i (ceux dans $F_n\backslash G_n$), on obtient :
$$\begin{aligned} |\{i \in F_n \mid c_U(T^i x)+i \notin F_n\}| &\leqslant |\{i \in G_n \mid c_U(T^i x) \notin K\}| + |F_n\backslash G_n| \\ &\leqslant |\{i \in F_n \mid c_U(T^i x) \notin K\}| + \frac{\delta}{2}|F_n| \\ &= \left(\sum_{i\in F_n} \mathbb{1}_{A_K}(T^i x)\right) + \frac{\delta}{2}|F_n| \end{aligned}$$
avec $A_K = \{y \in X \mid c_U(y) \notin K\}$. Par le théorème ergodique de von Neumann et l'inégalité de Bienaymé-Tchebychev, $X_n := \{x \in X \mid \frac{1}{|F_n|}\sum_{i\in F_n} \mathbb{1}_{A_k}(T^i x) \leqslant \mu(A_k) + \delta/4\}$ est de mesure $\geqslant 1-\delta$ pour n assez grand. Il suffit d'imposer que K soit assez grand de sorte que $\mu(A_K) \leqslant \delta/4$ et on obtient que pour n assez grand, pour tout x dans X_n (qui est de mesure $\geqslant 1-\delta$), on a $\sum_{i\in F_n} \mathbb{1}_{A_K}(T^i x) \leqslant (\delta/2)|F_n|$, donc au total,
$$|\{j \in \kappa_T(F_n,x) \mid 1+j \notin \kappa_T(F_n,x)\}| \leqslant \delta|F_n|.$$
La preuve est finie une fois qu'on a remarqué l'égalité $|F_n| = |\kappa_T(F_n,x)|$. □

C'est seulement dans la preuve du lemme 3 qu'on utilise le cocycle c_U, c'est-à-dire le fait que les U-orbites sont incluses dans les T-orbites (à conjugaison près).

Question 1. *Si on ne sait rien de l'inclusion des U-orbites dans les T-orbites, que peut-on dire ?*

Par exemple, si T vaut U^k, pour $k \geqslant 2$ et pour U d'entropie non nulle, on a $h(T) = kh(U)$ donc $h(T) > h(U)$. Cette inégalité a-t-elle toujours lieu lorsque le cocycle c_U n'existe pas ?

3.2 Preuve des lemmes 1 et 2

Preuve du lemme 1. ⚠ Dans toute cette preuve, si $\mathcal{P}$ est une partition et F une partie finie de $\mathbb{Z}$, alors $\mathcal{P}^F$ concernera l'action de T et pas celle de U : $\mathcal{P}^F = \bigvee_{m\in F} T^{-m}(\mathcal{P})$. Lorsqu'on fera référence à l'action de U, on utilisera la notation avec une observable : $(\mathcal{P} \circ U^m)_{m\in F}$.

Soit $k, n \in \mathbb{N}^*$, $I_n = \{0, k, 2k, \ldots, (n-1)k\}$, $R = [\![0, 1, \ldots, k-1]\!]$ et $F_n = R + I_n = [\![0, nk-1]\!]$ [8]. (F_n) est une suite de Følner. Considérons une partition finie $\mathcal{P}$ de X.

7. En première lecture, on peut considérer cette preuve "revisitée" dans le cas particulier où F_n est de la forme $[\![0,n]\!]$ (suffisant pour le lemme 1 car on le démontre avec la suite de Følner $F_n = [\![0, nk-1]\!]$). On prend $K = [\![-k,k]\!]$ avec k assez grand, et $G_n = [\![k, n-k]\!]$ pour n assez grand devant k. Si i est dans $[\![k,n-k]\!]$ et $c_U(T^i x)+i$ n'est pas dans $[\![0,n]\!]$, alors $c_U(T^i x)$ n'est pas dans $[\![-k,k]\!]$. Le terme d'erreur $|F_n\backslash G_n|$ vaut $2k$ qui est négligeable devant n.

8. Dans [6], on considère $F_n = RI_n$ où $I_n = \{e, g, \ldots, g^{n-1}\}$ avec $g \in G$ engendrant un sous-groupe distingué et d'indice fini, et R un système de représentant des classes modulo $\langle g\rangle$ (cf note 6 de bas de page sur la définition de $\widetilde{h}(\kappa_G)$ pour un groupe G virtuellement cyclique).

On veut faire apparaître U dans un majorant de $h(T,\mathcal{P})$. Pour cela, on va passer de l'action de T à l'action de U en utilisant la relation $T^{l}\cdot = U^{\kappa_T(l,\cdot)}\cdot$ et on va utiliser une partition telle que les exposants $\kappa_T(l,\cdot)$, pour certains l, soient constants sur chacune de ses pièces. Ainsi, conditionnellement à cette partition, on est en quelque sorte ramené au cas où T est une puissance de U. Cette partition est $(\mathcal{Q}_k^T)^{I_n}$, d'observable s'exprimant en fonction de $\kappa_T(k,\cdot)$.

$$h(T,\mathcal{P}) = \inf_{m\in\mathbb{N}} \frac{H(\mathcal{P}^{F_m})}{|F_m|} \leqslant \frac{H(\mathcal{P}^{F_n})}{|F_n|} \leqslant \frac{H(\mathcal{P}^{F_n} \mid (\mathcal{Q}_k^T)^{I_n})}{|F_n|} + \frac{H((\mathcal{Q}_k^T)^{I_n})}{|F_n|}.$$

- $\dfrac{H((\mathcal{Q}_k^T)^{I_n})}{|F_n|} \leqslant \dfrac{|I_n| H(\mathcal{Q}_k^T)}{|F_n|} = \dfrac{H(\mathcal{Q}_k^T)}{k}$.
- Pour $H(\mathcal{P}^{F_n} \mid (\mathcal{Q}_k^T)^{I_n})$, on va utiliser la notation avec les observables (cf sous-partie 2.2) pour plus de visibilité sur l'information contenue. On a $\mathcal{P}^{F_n} = \mathcal{P}^{R+I_n} = (\mathcal{P}^R)^{I_n} = \left(\mathcal{P}^R \circ T^{jk}\right)_{0\leqslant j\leqslant n-1}$ donc

$$H(\mathcal{P}^{F_n} \mid (\mathcal{Q}_k^T)^{I_n}) = \sum_{C\in(\mathcal{Q}_k^T)^{I_n}} \mu(C) H_{\mu_C}\left((\mathcal{P}^{F_n})_C\right) = \sum_{C\in(\mathcal{Q}_k^T)^{I_n}} \mu(C) H_{\mu_C}\left(\left(\mathcal{P}^R \circ T^{jk}{}_{|C}\right)_{0\leqslant j\leqslant n-1}\right)$$

Fixons $C \in (\mathcal{Q}_k^T)^{I_n}$. $\mathcal{Q}_k^T = \kappa_T(k,\cdot)$ donc $(\mathcal{Q}_k^T)^{I_n} = \left(\kappa_T(k, T^{ik}\cdot)\right)_{0\leqslant i\leqslant n-1}$ donc pour tout $0 \leqslant i \leqslant n-1$, $\kappa_T(k,T^{ik}\cdot)$ est constant sur C. En utilisant(1) avec $n=k$, on obtient que $\kappa_T(jk,\cdot)$ est constant sur C pour tout $0\leqslant j\leqslant n-1$ et donc que $T^{jk}{}_{|C} = U^{\kappa_T(jk,x_C)}{}_{|C}$ quelque soit x_C dans C, fixons un tel x_C. Au total, cela donne

$$\left(\mathcal{P}^R \circ T^{jk}{}_{|C}\right)_{0\leqslant j\leqslant n-1} = \left(\mathcal{P}^R \circ U^{\kappa_T(jk,x_C)}{}_{|C}\right)_{0\leqslant j\leqslant n-1} = \left(\mathcal{P}^R \circ U^m{}_{|C}\right)_{m\in\kappa_T(I_n,x_C)}$$

donc

$$H_{\mu_C}\left(\left(\mathcal{P}^R \circ T^{jk}{}_{|C}\right)_{0\leqslant j\leqslant n-1}\right) = H_{\mu_C}\left(\left(\mathcal{P}^R \circ U^m{}_{|C}\right)_{m\in\kappa_T(I_n,x_C)}\right) \leqslant \log\left|\left(\mathcal{P}^R \circ U^m\right)_{m\in\kappa_T(I_n,x_C)}\right|.$$

I_n est inclus dans F_n donc

$$\log\left|\left(\mathcal{P}^R \circ U^m\right)_{m\in\kappa_T(I_n,x_C)}\right| \leqslant \log\left|\left(\mathcal{P}^R \circ U^m\right)_{m\in\kappa_T(F_n,x_C)}\right|.$$

Au total,

$$\frac{H(\mathcal{P}^{F_n} \mid (\mathcal{Q}_k^T)^{I_n})}{|F_n|} \leqslant \sum_{C\in(\mathcal{Q}_k^T)^{I_n}} \mu(C) \frac{\log\left|\left(\mathcal{P}^R \circ U^m\right)_{m\in\kappa_T(F_n,x_C)}\right|}{|F_n|} = \sum_{C\in(\mathcal{Q}_k^T)^{I_n}} \mu(C) \frac{\log\left|\left(\mathcal{P}^R \circ U^m\right)_{m\in\kappa_T(F_n,x_C)}\right|}{|\kappa_T(F_n,x_C)|}.$$

Maintenant, supposons en plus que $\mathcal{P}$ est U-uniforme. Alors $\mathcal{P}^R$ est également U-uniforme. D'après la sous-partie 2.5, $\overset{\text{inv}}{\lim}_{F\to\infty} \frac{1}{|F|} \log\left|\left(\mathcal{P}^R \circ U^m\right)_{m\in F}\right|$ existe et est plus petite que $h(U)$. Soit $\varepsilon > 0$. On dispose de $\delta > 0$ tel que $\frac{1}{|F|}\log\left|\left(\mathcal{P}^R \circ U^m\right)_{m\in F}\right| \leqslant h(U) + \varepsilon$ pour tout F qui est $(\{1\},\delta)$-invariante. La $(\{1\},\delta')$-invariance implique la $(\{1\},\delta)$-invariance dès que $\delta' < \delta$ donc on peut supposer $\delta \leqslant \varepsilon$ sans perte de généralité.
D'après le lemme 3, pour n assez grand, on dispose de $X_n \subset X$ de mesure au moins $1-\delta$ et tel que pour tout $x \in X_n$, $\kappa_T(F_n,x)$ est $(\{1\},\delta)$-invariante [9].
Séparons les C qui s'intersectent avec X_n des autres. Soit $\mathcal{C}_n = \{C \in (\mathcal{Q}_k^T)^{I_n} \mid C \cap X_n \neq \varnothing\}$. Alors $\mu(\bigcup \mathcal{C}_n) \geqslant \mu(X_n) \geqslant 1-\delta \geqslant 1-\varepsilon$.
Pour les $C \notin \mathcal{C}_n$:

$$\sum_{C\in(\mathcal{Q}_k^T)^{I_n}\setminus\mathcal{C}_n} \mu(C) \frac{\log\left|\left(\mathcal{P}^R \circ U^m\right)_{m\in\kappa_T(F_n,x_C)}\right|}{|\kappa_T(F_n,x_C)|} \leqslant \sum_{C\in(\mathcal{Q}_k^T)^{I_n}\setminus\mathcal{C}_n} \mu(C) \frac{|\kappa_T(F_n,x_C)| \log|\mathcal{P}^R|}{|\kappa_T(F_n,x_C)|} \leqslant \varepsilon \log|\mathcal{P}^R|.$$

Pour les $C \in \mathcal{C}_n$:

9. La preuve "revisitée" du lemme 3 suffit ici puisqu'on y démontre l'énoncé pour la suite de Følner $[\![0,n]\!]$ dont (F_n) est une sous-suite.

On avait fixé x_C dans C, on va de plus imposer que x_C soit dans $C \cap X_n$, de sorte que pour n assez grand, les $\kappa_T(F_n, x_C)$ soient tous $(\{1\}, \delta)$-invariantes, ce qui donne :

$$\sum_{C \in \mathcal{C}_n} \mu(C) \frac{\log \left| \left(\mathcal{P}^R \circ U^m\right)_{m \in \kappa_T(F_n, x_C)} \right|}{|\kappa_T(F_n, x_C)|} \leqslant \sum_{C \in \mathcal{C}_n} \mu(C)(h(U) + \varepsilon) \leqslant h(U) + \varepsilon.$$

Au total, $h(T, \mathcal{P}) \leqslant \varepsilon \log |\mathcal{P}^R| + h(U) + \varepsilon + \frac{1}{k} H(\mathcal{Q}_k^T)$. En faisant tendre ε vers 0 et en prenant la borne inférieure sur $k \in \mathbb{N}^*$, cela donne $h(T, \mathcal{P}) \leqslant h(U) + \tilde{h}(\kappa_T)$. D'après la sous-partie 2.5, on a l'égalité $h(T) = \sup_{\mathcal{P}} h_\mu(T, \mathcal{P})$ où la borne supérieure ne porte que sur les partitions finies $\mathcal{P}$ qui sont U-uniformes. D'où le résultat.

$\square$

Preuve du lemme 2. Soit $k \in \mathbb{N}^*$. On veut un contrôle de $H(\mathcal{Q}_k^T)$ où $\mathcal{Q}_k^T$ admet pour observable $\kappa_T(k, \cdot)$. Intuitivement, $H(\mathcal{Q}_k^T)$ mesure l'incertitude de la valeur de $\kappa_T(k, x)$ pour x inconnu dans X. Or $\kappa_T(k, x) = c_T(x) + c_T(Tx) + \ldots + c_T(T^{k-1}x)$ donc l'incertitude de la valeur de $\kappa_T(k, x)$ dépend de l'incertitude de la valeur de chaque $c_T(T^i x)$. La stratégie est d'utiliser la bonne partition $\mathcal{P}_k$ pour que la majoration $H(\mathcal{Q}_k^T) \leqslant H(\mathcal{Q}_k^T \mid \mathcal{P}_k) + H(\mathcal{P}_k)$ ne soit pas trop grossière, au sens où la quantité $H(\mathcal{P}_k)/k$ ne soit pas trop grande et que sachant $\mathcal{P}_k$, on puisse lever un peu d'incertitude sur les $c_T(T^i x)$ pour que $H(\mathcal{Q}_k^T \mid \mathcal{P}_k)/k$ soit également petit. Nous allons trouver une telle partition conditionnellement à laquelle on sait que les $c_T(T^i x)$ prennent un nombre fini de valeurs et on s'autorisera à fixer la valeur de certains $c_T(T^i x)$ mais à titre exceptionnel pour que l'entropie de la partition ne soit pas trop grande.

Pour $r > 0$, considérons l'observable $f_r := c_T \mathbb{1}_{c_T \notin [\![-r,r]\!]}$. Ou bien elle donne la valeur exacte de $c_T(x)$ si $c_T(x)$ est hors de $[\![-r, r]\!]$, ou bien elle informe que $c_T(x)$ est dans $[\![-r, r]\!]$ sans donner la valeur exacte. La partition $\mathcal{P}_k$ que nous cherchons est celle associée à l'observable $\left(f_r \circ T^i\right)_{0 \leqslant i \leqslant k-1}$. Pour tout $0 \leqslant i \leqslant k-1$, ou bien elle donne la valeur exacte de $c_T(T^i x)$ si $c_T(T^i x)$ est hors de $[\![-r, r]\!]$, ou bien elle informe que $c_T(T^i x)$ est dans $[\![-r, r]\!]$ sans donner la valeur exacte.

Soit $\varepsilon > 0$. D'une part, pour r assez grand, on a $H(f_r) \leqslant \varepsilon$. Ainsi

$$H\left(\left(f_r \circ T^i\right)_{0 \leqslant i \leqslant k-1}\right) \leqslant k H\left(f_r\right) \leqslant \varepsilon k.$$

D'autre part, pour $D \in \left(f_r \circ T^i\right)_{0 \leqslant i \leqslant k-1}$, $H_{\mu_D}\left((\mathcal{Q}_k^T)_D\right) \leqslant \log \left|(\mathcal{Q}_k^T)_D\right| = \log \left|\kappa_T(k, \cdot)_{|D}\right|$. Par définition de $\left(f_r \circ T^i\right)_{0 \leqslant i \leqslant k-1}$, les $c_T \circ T^i{}_{|D}$ prennent une seule valeur ou sont dans $[\![-r, r]\!]$. Donc $\kappa_T(k, \cdot)_{|D}$ prend au plus $(2r+1)k$ valeurs [10]. Ainsi $H_{\mu_D}\left((\mathcal{Q}_k^T)_D\right) \leqslant \log\left((2r+1)k\right)$. Pour k assez grand, on a $\log\left((2r+1)k\right) \leqslant \varepsilon k$ donc

$$H\left(\mathcal{Q}_k^T \mid \left(f_r \circ T^i\right)_{0 \leqslant i \leqslant k-1}\right) = \sum_{D \in \left(f_r \circ T^i\right)_{0 \leqslant i \leqslant k-1}} \mu(D) H_{\mu_D}\left((\mathcal{Q}_k^T)_D\right) \leqslant k\varepsilon.$$

Au total, on a $H\left(\mathcal{Q}_k^T\right) \leqslant H\left(\mathcal{Q}_k^T \mid \left(f_r \circ T^i\right)_{0 \leqslant i \leqslant k-1}\right) + H\left(\left(f_r \circ T^i\right)_{0 \leqslant i \leqslant k-1}\right) \leqslant 2\varepsilon k.$

$\square$

Références

[1] T. Downarowicz. *Entropy in Dynamical Systems*. New Mathematical Monographs 18. Cambridge ; New York : Cambridge University Press, 2011.

[2] E. Glasner. *Ergodic Theory via Joinings*. T. 101. Mathematical Surveys and Monographs. Providence, Rhode Island : American Mathematical Society, 2003.

10. Dans le cas général d'un groupe H virtuellement abélien, cette borne sera plutôt de l'ordre de $e^{c|\Omega_D|}(rk)^j$ où j est l'ordre de croissance polynomiale du groupe et Ω_D est l'analogue (dans notre cas simple où H vaut $\mathbb{Z}$) de l'ensemble des i tels que $c_T \circ T^i{}_{|D}$ prend une seule valeur qui est hors de $[\![-r, r]\!]$. Lorsque qu'on passe au logarithme (cf la suite de la preuve), la puissance j ne modifie pas le raisonnement et l'exponentielle donne une quantité linéaire en $|\Omega_D|$. Il faut donc montrer que les D tels que $|\Omega_D| \geqslant \varepsilon k$ sont rares, pour cela on utilise principalement le théorème ergodique de von Neumann (on aura donc besoin de l'ergodicité de l'action de G qui est l'analogue de la transformation T) et l'inégalité de Bienaymé-Tchebychev.

[3] E. Glasner et B. Weiss. *On the Interplay between Measurable and Topological Dynamics*. 2004. arXiv : `math/0408328`.

[4] R. I. Jewett. « The prevalence of uniquely ergodic systems ». In : *Matematika* 17.4 (1973), p. 113-124.

[5] D. Kerr et H. Li. *Entropy, Shannon Orbit Equivalence, and Sparse Connectivity*. 2019. arXiv : `1912.02764 [math]`.

[6] D. Kerr et H. Li. *Entropy, Virtual Abelianness, and Shannon Orbit Equivalence*. 2022. arXiv : `2202.10795 [math]`.

[7] D. Kerr et H. Li. *Ergodic Theory : Independence and Dichotomies*. Springer Monographs in Mathematics. Springer International Publishing, 2017.

[8] W. Krieger. « On Unique Ergodicity ». In : *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2 : Probability Theory*. T. 6.2. University of California Press, 1972, p. 327-347.

[9] B. Weiss. « Strictly Ergodic Models for Dynamical Systems ». In : *Bulletin of the American Mathematical Society* 13.2 (1985), p. 143-147.